

ARIMA With GARCH Family Modeling and Projection on Share Volume of DSE

Ahammad Hossain

Varendra University, Rajshahi, Bangladesh

Md. Kamruzzaman, Md. Ayub Ali

University of Rajshahi, Rajshahi, Bangladesh

A suitable statistical model has been explored for the investors as well as the researchers to resolve the future estimation of share volume by using daily stock volume data from Dhaka Stock Exchange (DSE). The daily volume data from the June 1, 2004 to April 19, 2010 were retrieved from DSE website as a secondary data source. The Maximum Likelihood—Autoregressive Conditional Heteroskedasticity (ARCH) (Marquardt) method has been applied to construct the models for the stock volume data of DSE by using statistical package software E-Views of version-5. First of all, an “Auto Regressive Integrated Moving Average (ARIMA) model” was fitted and observed that heteroscedastic volatilities were still present there. To eliminate this dilemma, ARCH class of volatility models has been used and finally the ARIMA with EGARCH model has been explored. Findings of this study have recognized that ARIMA with EGARCH model implies low mean square error, low mean absolute error, low bias proportion, and low variance proportion for share volume data with comparing to other models. Hence, the modelling concept established in this study would be a decisive study for the investors as well as the researchers.

Keywords: ARIMA, Generalized ARCH (GARCH) family models, stock volume projection strategy

Introduction

Stock volume is a measure of market liquidity based on the number of shares that are traded over a given period. Volume data are recorded for individual stocks, their related options chains, and for indices as a whole. After price, volume is one of the most commonly quoted data points related to the stock market, reflecting the overall activity in a stock or market; the business of the market itself; and the buying and selling of shares (Chang, Cheng, & Pinegar, 1999). As such, volume is an important indicator for traders in analyzing market activity and planning strategy. Stock volume measures the changes in share prices that are generally associated with the changes in the market condition. The investors may regard it as a standard to determine the stock market condition in relation to earnings or dividend per share. Again the market condition of each company somehow depends on the economic condition of the country. Thus, stock volume gives an interpretation of how much stable the economic situation of a country. This study works with the daily share volume data of Dhaka

Ahammad Hossain, M.Sc., lecturer, Department of Natural Science, Varendra University, Rajshahi, Bangladesh.

Md. Kamruzzaman, Ph.D., associate professor, Institute of Bangladesh Studies, University of Rajshahi, Rajshahi, Bangladesh.

Md. Ayub Ali, Ph.D., professor, Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh.

Correspondence concerning this article should be addressed to Ahammad Hossain, Department of Natural Science, Varendra University, 529/1, Kazla, Motihar, Rajshahi-6204, Bangladesh. E-mail: ahammadstatru@gmail.com.

Stock Exchange (DSE) for searching suitable models. It may be one of the significant studies contributing to the unusual rise and fall in securities prices due to artificial manipulation of securities prices by a number of securities dealers and issuers in the absence of timely provision of reliable financial information in the market (Cootner, 1967). Firstly, the volume data have been explored and analyzed using visual inspection. Time series plot (in level, in first difference, in log transformation, in smoothing transformation, and combined of these) has been used to explore the data. The time series plot shows that the share volume of DSE has rightly upward trend over time but non-seasonality present in the series. The plot also indicates that volume series is heteroskedastic. But the log transformed time series plot violates the original trend and cannot remove the irregular variation of the data series. The differenced transformed series shows that the mean constant but the variance is not constant. The resistant smoothing method, 4253H-twice transformation (Velleman, 1980) series plot shows that there is a slightly upward trend over time and the irregular variations reduce from the data. Therefore, resistant smoothing transformed volume data series has been used in the whole analysis. To show the stationary condition of the series, the Dickey Fuller (DF) and Augmented Dickey Fuller (ADF) tests are used (Dickey & Fuller, 1979). The tests show the following results: The correlogram of the volume data series shows that the series is non-stationary. After non-seasonal differencing, the correlogram indicates that the series is stationary. DF and ADF tests indicate that the volume data series is non-stationary. After non-seasonal differencing of the series, both DF and ADF test suggest that the series is stationary. The ultimate goal of this econometric analysis is to find a suitable model that forecasts the share volume of DSE. Observing auto correlation function (ACF) and partial auto correlation function (PACF) (Box & Jenkins, 1976), an “Auto Regressive Integrated Moving Average (ARIMA) model” is fitted and observed that there are present heteroskedastic events. To select the best class of volatility model, different autoregressive conditional heteroskedasticity (ARCH) class models (Brooks, Faff, McKenzie, & Michell, 2000) are applied for comparative selection using Akaike information criteria (AIC), Bayesian information criterion (BIC) (Anderson & Burnham, 2002), and *R*-squared value. Finally, it is observed that the EGARCH models are quietly reasonable. Therefore, different types of ARIMA with EGARCH models have been used.

Review Literature of Modeling

The ARIMA provides features for the Box-Jenkins approach (Box & Jenkins, 1976) to the analysis of ARIMA models of univariate volume data series according to time. To adjust the unusual events of stock volume, ARIMA-intervention is used (Box & Tiao, 1975) with ARCH class volatility models (Abhyankar, 1995). Volatility models may be of two types as symmetric and asymmetric models (Bekaert & Wu, 2000). The main difference between the two classes is that symmetric models, including ARCH and generalized ARCH (GARCH), do not capture leverage effects in the time-series, as opposed to the asymmetric models. Asymmetric models include exponential GARCH (EGARCH) proposed by Nelson (1991), for details on asymmetric models see the work by Franses and Van Dijk (2000). Hillmer and Tiao (1982) used ARIMA technique for the seasonal adjustment as well as to introduce the decomposition the time series data into its mechanism, like trend, seasonal, and noise whereas such series will follow the assumption of the Gaussian ARIMA model. Engle (1982) proposed a model called ARCH model with the variation of conditional variance. In ARCH model, the restricted variance depends on the previous squared error terms of different lags, even at higher lag, one can grasp the maximum of the restricted variance, but a higher order indicates that the model comprises of several parameters which makes the estimation work lengthy, difficult, and different to intercept.

Later, Bollerslev (1986) proposed the GARCH model to conquer the higher order ARCH problem. The conditional variance depends on the previous squared errors and restricted variances of the GARCH model. The extension of ARCH through GARCH is like the extension of the AR to ARMA model, since the introduction of ARCH and GARCH models has been extensively used in this literature. Magnus and Fosu (2006) modeled and forecasted volatility by taking an individual index and using the models or specifications like GARCH (1, 1), EGARCH (1, 1), and TGARCH (1, 1). Rafique and Kashif-ur-Rehman (2011) studied the volatility clustering, excess kurtosis, and heavy tails of the time series using ARCH, GARCH, and Nelson's EGARCH processes (1991). It was found that GARCH (1, 1) has done the best to fully capture the persistence in volatility. The leverage effect was successfully overcome by EGARCH (1, 1) specification in KSE-100 index. Rodriguez and Ruiz (2009) studied the theoretical characteristics of a few and most trendy GARCH specifications having the component of leverage effect of positivity, stationarity, and finite fourth order moment. Floros (2008) used GARCH model and its subsequent variants for modeling and explaining volatility and financial market risk from daily observations from Egypt (CMA General Index) and Israel (TASE-100 index). Due to prices uncertainty during the time period under considerations, Egyptian CMA Index is the most volatile series. Tang, Chiu, and Leixu (2003) proposed the mixture of ARMA-GARCH model for volatile stock prices series and investigated its application in stock price prediction. In this paper, the finite mixture of ARIMA-GARCH (EGARCH) model has been formulated on stock volume series of DSE and investigated the forecasting performance of this model.

Empirical Methodology

ARIMA Model

Box and Jenkins (1976) theorized the ARIMA model. It has been formulized in the following way.

Differencing is often needed to make a time series stationary, supposing Y_t be a time series variable and considering the model:

$$Y_t - Y_{t-1} = C + \varepsilon_t$$

where, ε_t is a white noise disturbance term. Then, Y_t is said to be generated by a integrated process of order one and is denoted as I(1).

In a compact way, the model can be written as:

$$\nabla Y_t = C + \varepsilon_t$$

where, $\nabla = 1 - B$

Similarly, an integrated process of order p is denoted by I(d) and written as:

$$\nabla^d Y_t = C + \varepsilon_t$$

The theoretical ACF of an integrated process decays slowly (Barlett, 1946). From the above discussion, it is clear that an ARIMA process is nothing but the combination of the three processes: auto regression (AR) process, moving average (MA) process, and integrated process. The general ARIMA process of order p , d , and q is denoted by ARIMA (p , d , and q) and can be written in a compact way as follows:

$$\varphi(B)\nabla^d Y_t = C + \theta(B)\varepsilon_t \quad (1)$$

$\nabla^d = (1 - d)^d$ (The d order differencing operator)

$\varphi(B) = (1 - \varphi_1 B - \varphi_2 B^2 - \varphi_3 B^3 - \dots - \varphi_p B^p)$ (The p order AR operator)

$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_p B^p)$ (The q order MA operator)

where ε_t = random shocks, C is the constant, and Y_t is any time series.

When difference is not necessary to achieve stationarity, $d = 0$ and the model reduced to ARMA.

ARCH-Model

Over the past two decades, massive effort has been devoted to modeling and forecasting the movement of stock returns and other financial time series. Seminal work in this area of research can be recognized to Engle (1982) who introduced the standard ARCH model. Engle's process (1982) proposed to model time-varying conditional volatility using past innovations to estimate the variance using maximum likelihood estimation (MLE) (Bollerslev & Wooldridge, 1992) of the series as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (2)$$

where, ε_t denotes a discrete-time stochastic taking the form of $\varepsilon_t = z_t \sigma_t$ and $z_t \sim iid(0, 1)$; σ_t is the conditional standard deviation of return at time t .

GARCH-Model

Further extension introduced by Bollerslev (1986) is known as the GARCH model which suggests that the time-varying volatility process is a function of both past disturbances and past volatility. The GARCH model is an infinite order ARCH model generated by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_j \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_j \sigma_{t-j}^2 \quad (3)$$

where, α_0 , α , and β are non-negative constants. For the GARCH process to be defined, it is required that $\alpha > 0$.

EGARCH-Model

The EGARCH or exponential GARCH model was proposed by Nelson (1991). The specification for the conditional variance is:

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (4)$$

It should be noted that the left-hand side is of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic, and that forecasts of the conditional variance are guaranteed to be nonnegative. The presence of leverage effects can be tested by the hypothesis that $\gamma > 0$. The impact is asymmetric, if $\gamma \neq 0$. There are two differences between the E-Views specification of the EGARCH model and the original Nelson model (1991). First, Nelson (1991) assumed that the ε follows a generalized error distribution, while E-Views assumes normally distributed errors. Second, Nelson's specification (1991) for the log conditional variances differs slightly from the specification above:

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (5)$$

Estimating this model under the assumption of normal errors will yield identical estimates to those reported by E-Views except for the intercept term ω , which differ by $\alpha\sqrt{2/\pi}$.

$$\log \sigma_t^2 = \omega + \sum_{i=1}^p \beta_i \log \sigma_{t-i}^2 + \sum_{j=1}^q \left(\alpha_j \left| \frac{\varepsilon_{t-j}}{\sigma_{t-j}} - \sqrt{\frac{2}{\pi}} \right| + \gamma_j \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right) \quad (6)$$

The leverages effect term denoted as γ in the output, is negative and statistically different from zero, indicating the existence of the leverage effect in future bonds returns during the sample period.

Intervention Approaches

Generally, ARIMA process has been used to forecast a time series variable. But time series variable may sometimes be affected by some external events, for example, a labor strike may be found to coincide with a detected temporary decline in a product of a production firm; the production of rice may unexpectedly shift downwards as a cause of flood; the management of a company may take a policy for increasing the sales of their product. If these types of external events affect the series significantly, ARIMA process may not always be able to provide appropriate model to forecast the variable under study. Binary (0, 1) deterministic inputs can be used to represent such external events (Box & Tiao, 1975). The use of deterministic inputs in dynamic regression (DR) model to represent identified events is known as interventions. In time series analysis, interventions may be used to account for unexplained outliers, but the interests center on determining the uneven events that often occurred in time series variable. For a single intervention variable, authors use the rational form DR model as the framework to evaluate an intervention model.

The rational form DR model is given by

$$Y_t = C + \frac{w(B)B^b}{\delta(B)} I_t + E_t \quad (7)$$

where, Y_t is the time series variable under study; I_t is a binary deterministic variable; E_t is disturbance term described by an ARIMA process; and C is additive constant.

The Finite Mixture of ARIMA-EGARCH Model

The finite mixture of ARMA-GARCH model is similar to the mixture of AR-GARCH model proposed by Tang et al. (2003). Each component of the mixture can be denoted as a normal ARIMA series:

$$Y_t = C + \sum_{i=1}^I \theta_i Y_{t-i} + \sum_{j=1}^J \varphi_j \varepsilon_{t-j} + \sigma_t Z_t \quad (8)$$

$$\sigma_t^2 = K + \sum_{q=1}^Q A_q \sigma_{t-q}^2 + \sum_{p=1}^P G_p \varepsilon_{t-p}^2 \quad (9)$$

where, $k > 0$, $A_q \geq 0$, $G_p \geq 0$, and σ_t^2 are the conditional variance; Z_t is the standardized independent and identically distributed random variable drawn from some specified probability distribution; Z_t follows $N(0, 1)$ or student t -distribution with mean zero and variance unity and the degree of freedom parameter ν . Generalized expectation maximization (GEM) algorithm is used to learn the mixture model, authors can estimate tail quantities by assuming either the normal distribution or the t -distribution multiplying estimates of σ_t with the

standard quartiles of each distribution, and finally adding the conditional mean. Similarly for EGARCH model (6) can be expressed as:

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (10)$$

It should be noted that the left-hand side is of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic, and that forecasts of the conditional variance are guaranteed to be nonnegative. The presence of leverage effects can be tested by the hypothesis that $\gamma > 0$. The impact is asymmetric, if $\gamma \neq 0$.

Model Selection Criteria

Mainly mean squared error (MSE) and mean absolute error (MAE) have been used for model selection in this analysis. The mathematical formula of MSE and MAE are as follows:

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$$

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$$

There are some other statistics for model selection. They are AIC and BIC. Authors choose models that give smallest value of these criteria. AIC (Sakamoto, Ishiguro, & Kitagawa, 1986) is one of the most important criteria for checking the adequacy as well as the lag order of a model. AIC is defined as:

$$\text{AIC} = \log \left(\frac{\sum \hat{\varepsilon}_i^2}{N} \right) + \frac{2k}{N}$$

where $\sum \hat{\varepsilon}_i^2$ is the sum of squared residuals.

In principle, one could select a lag structure by increasing the number of lags up to the point where the AIC reaches a minimum value. The BIC is another penalized maximum likelihood criteria and was originally proposed by Schwarz in 1978. BIC was derived in a Bayesian context and approximates a variant of Laplace's method (Lindley, 1980). The BIC criterion is:

$$\text{BIC} = \log(\hat{\sigma}^2) + \frac{n \log(N)}{N}$$

Where, n is the dimensionality of the model; $(1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) y_t = c + (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t$ is the estimated of the variance; and N is the sample size.

Projecting Algorithm Technique

This is a very difficult question to choose the best algorithm. Since, real data do not follow any model.

Some general advices are as follows. Firstly, to identify what measure of forecast error is most appropriate for the particular situation at hand (Pankratz, 1991).

(1) MSE, MAE, AIC, BIC, etc. have been used for appropriate model selection;

(2) Assuming enough historical data is available, and then there proceed as follows: The variety of algorithms and projecting techniques have been used to predict the next k observations for the training set and comparing the projection to the actual realized values (the test set). Using this projecting technique that gives the smallest value of MSE for the test set on the original data set (training set) is to obtain the desired out sample forecasts or projected values.

To assess the performance of the finite mixture of ARIMA-EGARCH models, the following measures of statistic fit are compared:

- (1) MSE;
- (2) MAE;
- (3) Bias Proportion (BP);
- (4) Variance Proportion (VP).

Empirical Results and Discussion

At first, the volume data of DSE have been explored and analyzed using visual inspection. Time series plot (in level, in first difference, in log transformation, in smoothing transformation, and combined of these) has been used to explore the necessary information of stock volume series (Ding, Granger, & Engle, 1993). Plots have shown the following results: The time series plot (Granger & Andersen, 1978) shows that the share volume of DSE has rightly upward trend over time but non-seasonality present in the series; there are heteroskedastic events present in the series (Christofi & Pericli, 1999). But the log transformed time series plot violates the original trend and cannot remove the irregular variation of the data series. The differenced transformed series shows that the mean constant but the variance is not constant. The resistant smoothing method, 4253H-twice transformation series plot shows that there is slightly upward trend over time and the irregular variations reduce from the data (Figure 1). Therefore, smoothing transformed volume data series has been used in the whole analysis.

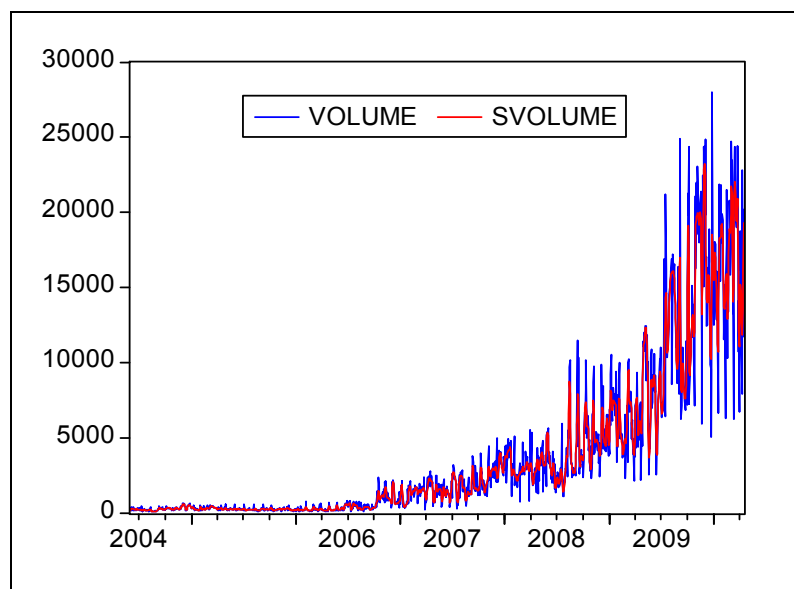


Figure 1. Combined time series plot between daily volume data and smoothing transformed volume data of DSE.

To show the stationary condition of the series, the DF and ADF tests are used there. DF and ADF tests suggest that the non-differenced volume data series is insignificant and differenced transformed volume data series is highly significant at 5% and 10% level of significance. Therefore, the volume data series is non-stationary and after non-seasonal differencing of the series, both DF and ADF tests suggest that the series is stationary (Table 1).

Table 1

DF and ADF Test of Volume Data Series (Y_t) and Differenced Transformed Volume Data Series (ΔY_t)

Data	Test	Test statistic (Prob.)	Asymptotic critical values at 5%	Asymptotic critical values at 10%
Y_t	DF	0.936619 (0.891)		
	ADF	0.210989 (0.863)		
ΔY_t	DF	-7.087286 (0.000)	-2.863242	-2.567724
	ADF	-11.28614 (0.000)		

The ultimate goal of this study is to find a suitable model to project the future trend of share volume of DSE. Observing ACF and PACF, an ARIMA (1, 1, 1) model is fitted (Figure 2) and it is observed that there are still present heteroskedastic events.

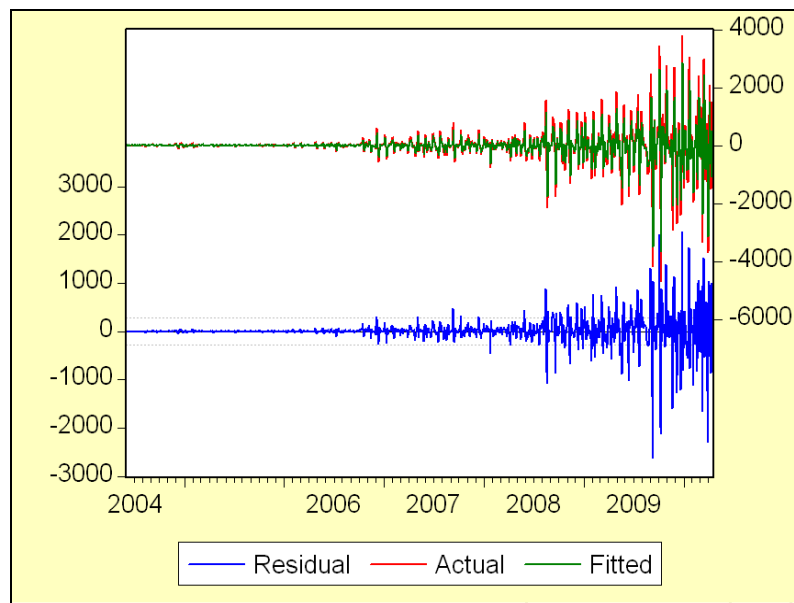


Figure 2. Actual, fitted, and residual plot obtained from ARIMA (1, 1, 1) model.

The residual of fitted an ARIMA (1, 1, 1) model was stationary, but the residuals in some points are varying abnormally, suggesting that unusual belongings or heteroskedastic or intervention shocks are present there. The coefficients ARIMA (1, 1, 1) model is not consistent. The R -squared value of ARIMA (1, 1, 1) model is 0.856416, which means that about 85% of the variation of daily share volume data are explained by the model. So, ARIMA (1, 1, 1) model may not be a suitable fitted model. The estimation result of ARIMA (1, 1, 1) model is shown in Table 2.

To select the best class of volatility model, different ARCH class models (Chand, Kamal, & Ali, 2012) are used for comparative selection using AIC, BIC, and *R*-squared value (Table 3). Finally, it is observed that the EGARCH models are quietly reasonable.

Table 2

Summary Statistics of Model ARIMA (1, 1, 1)

Variable	Coefficient	Std. error	T-statistic	P-value. *(less than 0.05)
AR (1)	1.618158	0.024426	66.24823	0.0000
MA (1)	0.317581	0.024590	12.91528	0.0000

Note. * All the coefficients are highly significant at 5% level of significance.

Table 3

Comparative Selecting Table for the Best Fitted ARCH Class Volatility Models

Models	Coefficients (Prob.)					AIC	BIC	R-squared value
ARCH (1)	95,080.6 *(0.000)	0.43687 *(0.0000)				13.6650	13.6894	0.88656
ARCH (2)	63,771.59 *(0.000)	0.30548 *(0.0003)	0.25427 *(0.000)			13.3554	13.3833	0.88356
ARCH (3)	244,168.1 *(0.000)	-0.01311 *(0.7798)	0.18326 (0.0509)	-0.08487 *(0.0180)		14.4081	14.4396	0.88625
GARCH (1, 1)	226,956.8 *(0.000)	0.024008 *(0.000)	-0.99992 *(0.000)			13.35010	13.37804	0.873216
GARCH (2, 1)	225,004.0 *(0.000)	0.001916 (0.1912)	-0.94563 *(0.000)	-0.94870 *(0.000)		13.59199	13.62343	0.887988
GARCH (2, 2)	232,132.0 *(0.000)	-0.00045 (0.6592)	0.000795 (0.2595)	-0.99718 *(0.000)	-0.9972 *(0.000)	13.21024	13.24517	0.887984
TARCH (1, 1, 0)	229,420.0 *(0.000)	0.011252 (0.8712)	0.123283 (0.2944)			14.36146	14.38940	0.887140
TARCH (1, 1, 1)	226,831.6 *(0.000)	0.040521 *(0.000)	-0.03941 *(0.000)	-0.99999 *(0.000)		12.53511	12.53511	0.884161
* EGARCH (1, 1, 1)	-0.22143 *(0.000)	0.389713 *(0.000)	0.067552 *(0.000)	0.994519 *(0.000)		10.58747	10.61891	0.883349
* EGARCH (2, 1, 1)	0.284980 *(0.000)	0.581958 *(0.000)	0.070497 *(0.0003)	0.439019 *(0.000)	0.546847 *(0.000)	10.56697	10.60190	0.885952

Notes. The (*) marked indicated Prob. value represents that the estimated coefficient models are significant and the rest of them are insignificant at 5% level of significance.

There is also found that EGARCH (1, 1, 1) and EGARCH (2, 1, 1) models have the lower value of AIC and BIC value (model choosing criteria) and the higher value of *R*-squared value (coefficient of determination) what is expected. Therefore, ARIMA with exponential GARCH models have been used for volume data series of DSE. Therefore, three ARIMA with EGARCH models are constructed, but one of them (an intervention module) is applied. The results of diagnostic checking, model comparison, model selection, and projecting performance of these models are summarized as follows: The estimated values, standard errors, Z-statistic, and *p*-values of model 1, model 2, and model 3 are reported in tables (Table 4, Table 5, and Table 6). The *p*-values showed that the entire coefficient is highly significant. The residual standard error of the shock term series is unknown; sample standard error is used to estimate it. The residual standard error for model 1 is 207.9113, model 2 is 288.6837, and model 3 is 200.1666. The *R*-squared value for model 1 is 0.876470, model 2 is 0.754143, and model 3 is 0.885952 respectively.

The below three models have an interesting interpretation. For every model, the absolute values of the parameters are less than unity, which ensures the stationarity condition of the models (Ding & Granger, 1996). The R -squared value for model 1 is 0.876470 meaning that 87% of the variation in the daily volume data is explained by model 1; the R -squared value for model 2 is 0.754143 meaning that 75% of the variation in the daily volume data is explained by model 2, and the R -squared value for model 3 is 0.885952 meaning that 88% of the variation in the daily volume data is explained by model 3. Therefore, the R -squared value obtained from model 3 is greater than from model 1 and model 2. Therefore, model 3 should be the better fitted model. The observed R -squared values for the model 1, model 2, and model 3 are 0.875658, 0.752688, and 0.885275. The adjusted R -squared values suggest that the sample regression line fit the data well. Actual, fitted and residual plots obtained from model 1, model 2, and model 3 have been shown in figures (Figure 3, Figure 4, and Figure 5). It is observed that the model 1 and model 2 do not so well fit; model 3 is fine fitted, but at the end, there is a little fluctuation. Therefore, there may conclude that the model 3 may be the appropriate fitted model.

Table 4

Summary of Model 1

ARIMA (1, 1, 4)	Coefficient	Std. error	Z-statistic	Prob.
C	-0.626521	0.362007	-1.730690	0.0835
AR (1)	0.825675	0.001087	759.6219	0.0000 *
MA (1)	0.870679	0.000541	1,608.938	0.0000 *
MA (2)	-0.364910	0.000254	-1,436.819	0.0000 *
MA (3)	-0.612601	0.001354	-452.2973	0.0000 *
MA (4)	-0.385257	0.000982	-392.5003	0.0000 *
EGARCH (2, 1, 1)				
ω	-0.273765	0.011534	-23.73468	0.0000 *
α	0.401401	0.018661	21.50991	0.0000 *
γ	0.225513	0.015978	14.11419	0.0000 *
β_1	0.328542	0.038799	8.467850	0.0000 *
β_2	0.671079	0.038818	17.28804	0.0000 *

Note. All the (*) marked coefficients are significant at 5% level of significance.

Table 5

Summary of Model 2

ARIMA-intervention	Coefficient	Std. error	Z-statistic	Prob.
I_1	-543.7147	100.4694	-5.411744	0.0000
AR (1)	0.955598	0.005835	163.7597	0.0000
MA (2)	-0.385246	0.027325	-14.09889	0.0000
MA (3)	-0.270217	0.032678	-8.269200	0.0000
MA (4)	-0.320691	0.024840	-12.91015	0.0000
EGARCH (2, 1, 1)				
ω	-0.530910	0.026081	-20.35595	0.0000
α	0.887930	0.025879	34.31145	0.0000
γ	0.119843	0.022390	5.352584	0.0000
β_1	0.272405	0.021037	12.94883	0.0000
β_2	0.714189	0.020721	34.46721	0.0000

Notes. All the coefficients are significant at 5% level of significance and here; I_1 represents the intervention component in ARIMA process.

Table 6

Summary of Model 3

ARIMA (8, 1, 1)	Coefficient	Std. error	Z-statistic	Prob.
AR (1)	0.708849	0.017663	40.13153	0.0000
AR (4)	-0.430449	0.024146	-17.82715	0.0000
AR (5)	0.181059	0.028515	6.349677	0.0000
AR (8)	-0.019781	0.011498	-1.720390	0.0854
MA (1)	0.997513	0.000965	1,033.623	0.0000
EGARCH (2, 1, 1)				
ω	-0.284980	0.017134	-16.63198	0.0000
α	0.581958	0.024898	23.37326	0.0000
γ	0.070497	0.019559	3.604394	0.0003
β_1	0.439019	0.047384	9.265160	0.0000
β_2	0.546847	0.046988	11.63799	0.0000

Note. All the coefficients are significant at 5% level of significance except only the AR (8).

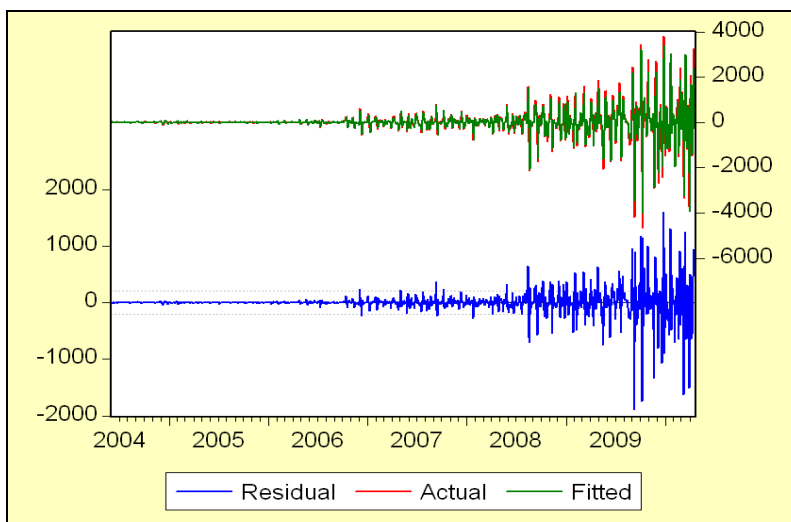


Figure 3. Actual, fitted, and residual plot obtained from model 1.

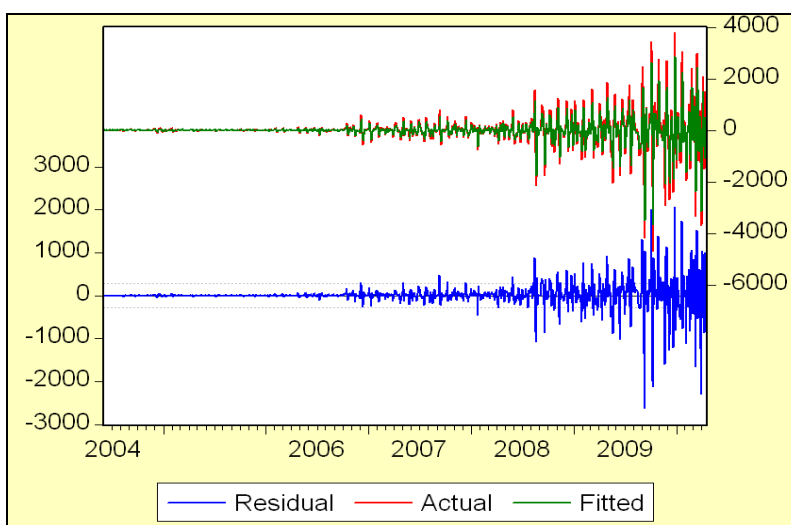


Figure 4. Actual, fitted, and residual plot obtained from model 2.

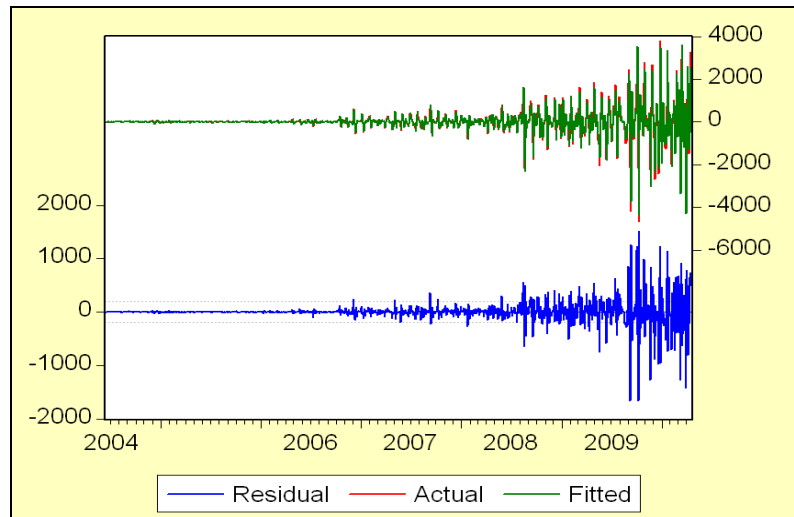


Figure 5. Actual, fitted, and residual plot obtained from model 3.

The value of the Jarque Bera test statistic for the model 1 is 1,233.776 with significant probability (P -value = 0.000); model 2 is 453.1224 with significant probability (P -value = 0.000); and model 3 is 1,153.019 with significant probability (P -value = 0.000). So, it may conclude that the residual obtained from these models may be normal. The forecasting evaluation, model selection criteria, and finally, the linear trend of forecasting performance of the proposed model have been checked. MSE, MAE, BP, and VP of model 3 are minimum than model 1 and model 2. Therefore, model 3 should be the most adequate model to project the volume data series of DSE (Table 7).

Table 7

Results of Forecasting Evaluation of Model 1, Model 2, and Model 3

Forecast sample: December 15, 2009 to April 19, 2010			
Measures of statistic fit	Model 1	Model 2	Model 3
MSE	207.1641	287.5877	199.5095
MAE	65.17870	128.6973	87.39071
BP	0.078875	0.08945	0.000361
VP	0.817367	0.897231	0.002790

The AIC and BIC of model 3 are minimum (Table 8). Therefore, the model 3 has been selected.

Table 8

Results of AIC and BIC for the Model 1, Model 2, and Model 3

Model	AIC	BIC
Model 1	10.58943	10.62811
Model 2	11.063116	11.09800
Model 3	10.56697	10.60190

Forecast values generated (test series) by model 3 are higher than the observed values (actual or training series), but observed values and projecting (Forecast) values are approximately identical and the forecast values have the linear upward trend (Figure 6). Since the forecast values are approximately identical to observed values, the out sample projecting performance of model 3 is quietly reasonable.

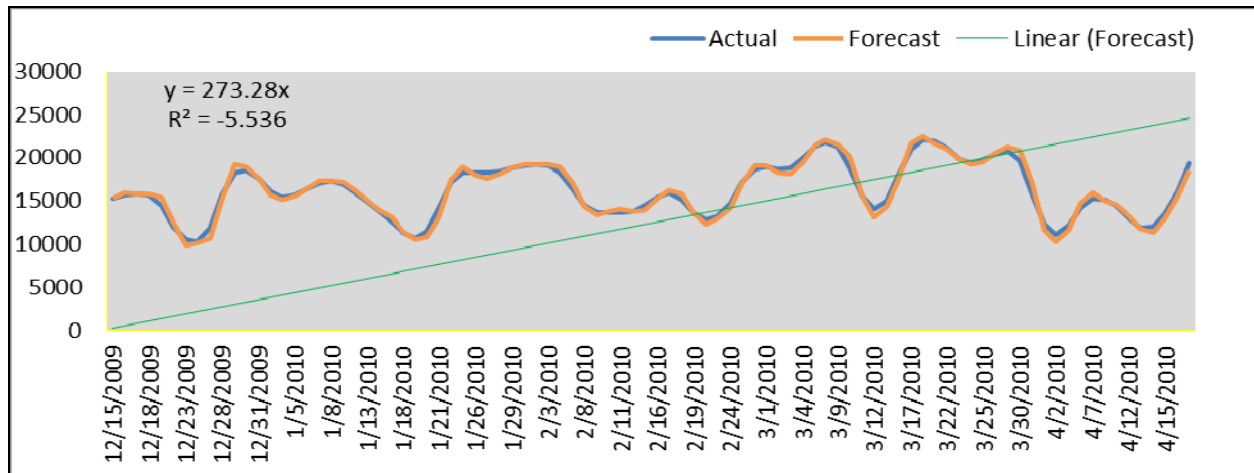


Figure 6. Out sample forecasts of last 90 observations from the selected model.

Conclusions

Daily share price of DSE Ltd. is very blazing concern in Bangladesh and it is necessary for policy implications. The result of this study may help Bangladesh Securities and Exchange Commission (BSEC) and the government of Bangladesh to take appropriate actions to ensure the stock market condition. Share volume deals with the changes in share prices that are usually associated with the changes in the market condition. The findings established that ARIMA with EGARCH model comprises low MSE, low MAE, low BP, and low VP for volume data and thus, the modeling concept established in this paper would be useful for the investors or researchers to determine the future trend of share volume and there by taking decision for investment.

References

- Abhyankar, A. H. (1995). Trading-round-the clock return, volatility and volume spillovers in the Eurodollar futures markets. *Pacific-Basin Finance Journal*, 3, 75-92.
- Anderson, D. R., & Burnham, K. P. (2002). Avoiding pitfalls when using information-theoretic methods. *Journal of Wildlife Management*, 66, 910-916.
- Barlett, M. S. (1946). On the theoretical specification of sample properties of auto correlated time series. *Journal of the Royal Statistical Society*, 27, 27-41.
- Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13, 1-42.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-328.
- Bollerslev, T., & Wooldridge, J. (1992). Quasi-maximum likelihood estimation inference in dynamic models with time-varying covariances. *Econometric Theory*, 11, 143-172.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Applied times series analysis and Box-Jenkins models*. Orlando: Academic Press.
- Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environment problems. *Journal of the American Statistical Association*, 70, 349.
- Brooks, R. D., Faff, R. W., McKenzie, M. D., & Michell, H. (2000). A multi-country study of power ARCH models and national stock market returns. *Journal of International Money and Finance*, 19, 377-397.
- Chand, S., Kamal, S., & Ali, I. (2012). Modeling and volatility analysis of share prices using ARCH and GARCH models. *World Applied Sciences Journal*, 19(1), 77-82.
- Chang, E. C., Cheng, J. W., & Pinegar, J. (1999). Does futures trading increase stock market volatility? The case of the Nikkei stock index futures markets. *Journal of Banking and Finance*, 23, 727-753.
- Christofi, A., & Pericli, A. (1999). Correlation in price changes and volatility of major Latin American stock markets. *Journal of Multinational Financial Management*, 9, 79-93.

- Cootner, P. H. (1967). *The random character of stock market prices*. Cambridge: The M. I. T. Press.
- Dickey, D. A., & Fuller, W. A. (1979). Distributions of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association*, 74, 427-481.
- Ding, Z., & Granger, C. W. J. (1996). Modeling instability persistence of speculative returns: A new approach. *Journal of Econometrics*, 73, 185-215.
- Ding, Z., Granger, C. W. J., & Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, 83-106.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, 50, 987-1008.
- Floros, C. (2008). Modeling volatility using GARCH models: Evidence from Egypt and Israel. *Middle Eastern Finance and Economics*, 2, 31-41.
- Franses, P. H., & Van Dijk, D. (2000). *Nonlinear time series models in empirical finance*. New York: Cambridge University Press.
- Granger, C. W. J., & Andersen, A. P. (1978). *Introduction to bilinear time series models*. Göttingen: Vandenhoeck and Ruprecht.
- Hillmer, S. C., & Tiao, G. C. (1982). An ARIMA-model based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77, 63-70.
- Lindley, D. V. (1980). Approximate Bayesian methods. *Trabajos de Estadística Y de Investigación Operativa*, 31, 223-245.
- Magnus, F. J., & Fosu, O. (2006). Modeling and forecasting volatility of on returns on the Ghana stock exchange using GARCH models. *Journal of Applied Science*, 3, 2042-2048.
- Nelson, D. B. (1991). Conditional heteroskedasticity in assets returns: A new approach. *Econometrica*, 59, 347-370.
- Pankratz, A. (1991). *Forecasting with dynamic regression models*. New York: Johan Wiley & Sons Inc.
- Rafique, A., & Kashif-ur-Rehman. (2011). Comparing the persistency of different frequencies of stock returns volatility in an emerging market: A case study of Pakistan. *African Journal of Business Management*, 5, 59-67.
- Rodriguez, M. J., & Ruiz, E. (2009). *GARCH models with leverage effect: Difference and similarities* (Working Paper, Statistics and Econometrics Series, 02 January 2009, Departamento de Estadística Universidad, Spain).
- Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. Dordrecht: D. Reidel Publishing Company.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464.
- Tang, H., Chiu, K. C., & Xu, L. (2003). Finite mixture of ARMA-GARCH model for stock price prediction. Proceedings from *3rd International Workshop on Computational Intelligence in Economics and Finance*, North Carolina, USA.
- Velleman, P. F. (1980). Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association*, 75, 609-615.